Predicting Improvement of Functioning in Disability Claimants

Nieuwenhuijsen, K.; Cornelius, L. R.; de Boer, M.R.; Groothoff, J. W.; Frings-Dresen, M. H. W.; van der Klink, J. J. L.; Brouwer, Sandra

Published in:
Journal of Occupational Rehabilitation

DOI:
10.1007/s10926-013-9469-2

J Occup Rehabil (2014) 24:410–418
DOI 10.1007/s10926-013-9469-2

Predicting Improvement of Functioning in Disability Claimants

K. Nieuwenhuijsen · L. R. Cornelius · M. R. de Boer · J. W. Groothoff · M. H. W. Frings-Dresen · J. J. L. van der Klink · S. Brouwer

Published online: 12 September 2013

Abstract *Purpose* In the Netherlands, disability claimants are assessed after 2 years of sick leave, but their functioning may still improve. An accurate prognosis of functioning is difficult. Self predictions may be more accurate than those of professionals. The aim of this study, is to assess and compare the accuracy of predictions by disability claimants and insurance physicians (IPs) working at the Social Security Institute. It is further studied whether the accuracy differs between subgroups of claimants with mental or somatic health conditions. *Methods* We used data from the prospective cohort study cohort study. Following the assessment of the disability claim (n = 375) and after 1 year follow up (T1, n = 276) data on functioning were obtained from respondents by self-report questionnaire World Health Organization Disability Schedule 2.0. Both claimants and IPs were asked to predict improvement of functioning. Accuracy of their predictions were assessed by sensitivity, specificity, and area under the receiver operating curves (AUC). Mixed logistic regression was conducted to explore differences in accuracy between claimants with mental and somatic conditions. *Results* One-third (32 %) of disability claimants improved beyond the standard error of measurement. Disability claimants' and IPs were able to predict this improvement of functioning, but to a limited extent, with an AUC of 0.61 for IPs and 0.62 for disability claimants. We found no statistically significant differences in the accuracy of the predictions in claimants with mental or somatic health conditions. *Conclusions* Improvements of functioning were not uncommon. However, both IPs and disability claimants were unable to predict improvement with high levels of accuracy in both mental and somatic health conditions.

Keywords Disability evaluation · Sick leave · Occupational health · Insurance claim review

K. Nieuwenhuijsen (✉) · M. H. W. Frings-Dresen
Coronel Institute of Occupational Health, Academic Medical Center (AMC), University of Amsterdam, P.O. Box 22700, 1100 DE Amsterdam, The Netherlands
e-mail: k.nieuwenhuijsen@amc.nl

K. Nieuwenhuijsen · M. H. W. Frings-Dresen
Research Center for Insurance Medicine, Amsterdam, The Netherlands

L. R. Cornelius · M. R. de Boer · J. J. L. van der Klink · S. Brouwer
Research Center for Insurance Medicine, Groningen, The Netherlands

L. R. Cornelius · J. W. Groothoff · J. J. L. van der Klink · S. Brouwer
Department of Health Sciences, Community and Occupational Medicine, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

L. R. Cornelius
Social Security Institute, Amsterdam, The Netherlands

M. R. de Boer
Department of Health Sciences, VU University, Amsterdam, The Netherlands

Introduction

Long term sickness absence and work disability have a major impact on most industrialized societies. In 2007 the Organization for Economic Co-operation and Development (OECD) calculated that in OECD countries an average of 5.8 % of the working population received sickness absence or disability benefits. These benefits amounted to an average of 1.2 % of their gross domestic product [1]. These numbers stem from before the economic downfall. Therefore, based on past experiences, the OECD expects an increase of disability beneficiaries in these economic circumstances [2].

Apart from societal costs, long term sickness absence and work disability have negative consequences for affected workers. Workers with various mental and somatic disorders emphasize the importance of work to their wellbeing [3–5]. Consequences of not being able to work include lower mental health and quality of life [6, 7]. Sickness absence and work disability in workers with various specific or any health conditions have even been shown to increase the risk of mortality compared to those able to stay at work despite having these conditions [8–10].

Disability benefits serve the important purpose of providing financial security for persons unable to work for an extended period of time. Many of the health conditions for which disability benefits are granted may show improvements or deteriorations over time. Such changes have been observed in for instance depressive disorders [11] and low back pain [12]. While being granted disability benefit may meet the recipients' need to be financially secure, ongoing disability while improvements may occur may have negative consequences. In some countries disability benefits do not require reassessments [13]. A lack of resources affects the possibility to regularly re-assess eligibility in other countries [14]. Therefore, it is vital to identify those disability claimants whose functioning may still improve in the future. These benefit recipients may be more closely monitored in order to appropriately time return-to-work interventions.

To this aim, an accurate prognosis of health status and functioning of disability claimants is needed. However, predicting future improvement of functioning is difficult. In workers on sick leave for any diagnosis, general physicians were able to predict the absence status only 4 weeks later accurately in 53 % of the cases with sick leave durations of 3–20 weeks [15]. One of the reasons for this may be that health-related disability appears to be a bio-psycho-social phenomenon [16]. Other factors than the medical status also influence the prognosis of disability [17]. Several studies have shown that the return-to-work perceptions of workers on sick leave are related to the actual return to work. In studies of workers with common mental disorders it was found that return-to-work perceptions, i.c. predicted duration of sickness absence or likelihood of return to work, is a predictor of future actual return to work [18–21]. Similar results were found in studies in workers with back pain with regard to return-to-work perceptions, in this case the perceived ability of the worker to meet work demands after returning [22, 23] and the duration of sickness absence benefits predicted by the worker [24, 25]. Moreover, in workers on sick leave due to musculoskeletal or mental disorders, self-predictions more accurately predicted the length of sick leave than professionals who based the prediction on information in the medical certificates [26]. Finally, Wind et al. [27] concluded that disability claimants were able to predict the outcome of their application for a disability benefit.

Perceptions of disability claimants about the likelihood and time of future improvements may be a valuable source of information for the disability assessment. Perceptions about improvements of functioning may be predictive of actual improvements. If so, the workers' perspective may be an essential part of the disability benefit assessment. In the Netherlands, social insurance physicians (IPs) are responsible for the disability benefit assessment in cooperation with labor experts. The assessment takes place if the employee has not fully returned to work after 2 years of sickness absence. Disability benefits are granted when a worker has a substantial loss of his or her earning capacity due to health reasons. The benefit is granted regardless of the work-relatedness of the health condition. Benefit reassessments are scheduled according to the expected improvements in a recipients health or functioning.

Self-perceptions have been hypothesized to be associated with future work outcomes but studies have yielded inconsistent results for workers with somatic and mental health conditions. Workers on sick leave with somatic health conditions did report higher levels of self-efficacy to return to work compared to those with mental health problems [28]. Also, differential associations between self-efficacy and actual return to work were observed over health conditions. Moreover, self-efficacy was associated with impairment and disability in welfare claimants with mental health conditions only [29]. In contrast, perceived time to return to work was related to actual return to work across health conditions in another study of workers on non-work-related sick leave [30]. With regard to the perceptions of professionals, one study found that the probability of doctors making accurate predictions differed across health conditions [15].

The objective of this study is therefore to assess the diagnostic accuracy of the prediction of improvement of functioning by disability claimants and to compare it with that of the IPs assessing their disability claim. Furthermore, this study aims to investigate whether the diagnostic accuracy differs between subgroups of claimants with mental or somatic health conditions.

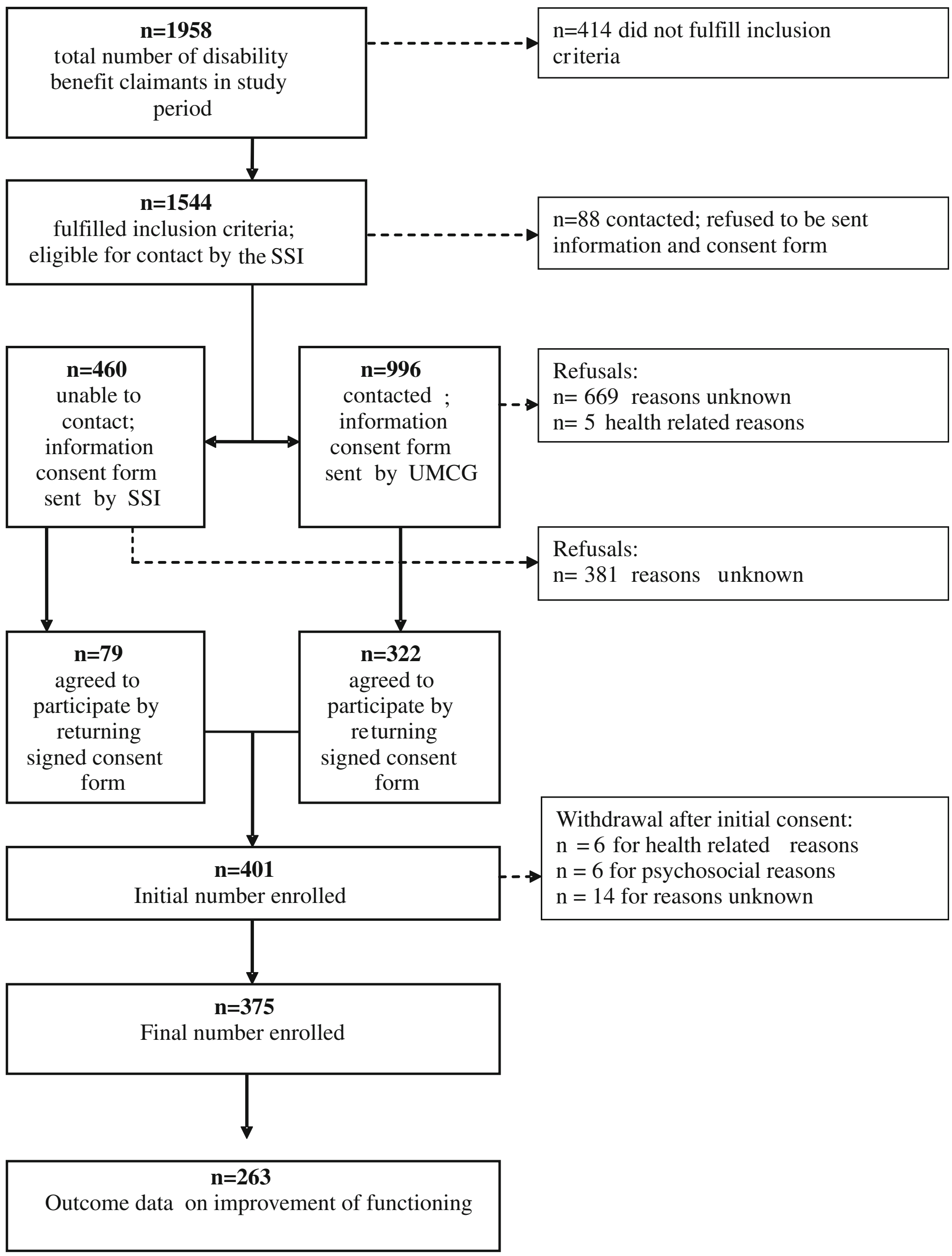
Methods

Participants and Procedures

In the Dutch social security system disability claims are assessed by IPs and labour experts employed by the Dutch Social Security Institute (SSI). For the present study, data were drawn from a prospective cohort study (PREDIS) with 1 year follow-up among persons claiming disability benefit after 2 years of sickness absence. All diagnoses were included, both mental and somatic. Eligible participants were recruited using registry data from the local SSI office in the city of Groningen, servicing Groningen and Drenthe, two northern provinces of the Netherlands. Recruitment started at October 1st 2008 and ended at 31st December 2009. Follow up was conducted independently from outcome of the disability claim. Follow up ended at April 1st 2011. The Medical Ethics Committee of the University Medical Center Groningen, the Netherlands, approved recruitment, consent and field procedures. Out of a total of 1,544 eligible disability claimants, 375 persons consented to participate, adding up to a response rate of 24.3 %, see Fig. 1.

To assess the representativeness of the study sample (n = 375) for the national population of persons claiming disability benefit in the Netherlands, we used data from the SSI on demographic characteristics [31] and diagnostics [32]. We found the sample slightly different to the national population as to the prevalence of mental disorders (study sample 22 % and national population 34 % as certified by the SSI as primary cause of disability in 2009).

Fig. 1 Flowchart of participants



Design

At baseline following the assessment of the disability claim (T0), and after 1 year follow up (T1) data on functioning were obtained from respondents by self-report questionnaire. Diagnostic data certified as cause of disability at T0 were obtained from IPs after their assessment of the disability claim. Both respondents and the IPs assessing their disability benefit claims were asked to predict improvement of functioning.

Measures

Demographic Characteristics

Data on age and gender were obtained during a structured clinical interview for classifying mental disorders (CIDI [33]) that was part of the PREDIS study. Data on educational level were obtained from the SSI labor experts assessing the disability claim and data on work status were provided by the SSI register.

Diagnosis

To classify medical diagnoses, IPs use a classification system (Dutch Classification for Occupational Health and Social Insurance: CAS) derived from the ICD-10 [34] and developed for use in occupational health and social security in the Netherlands [35]. From the IPs we obtained the ICD-10 codes of the somatic and mental disorders certified as the primary cause for the claimants disability.

Prediction of Functioning by Social Insurance Physician At baseline, the prediction of improvement in functioning by IPs was assessed by asking whether they expected improvements in functioning (yes, no, unsure) and if so, at what time they expected this improvement to occur (in months). Answers were dichotomized as expecting an improvement of functioning within 1 year versus expecting an improvement later than 1 year or not at all. Being unsure about improvement was not included in this dichotomous variable. The IPs answered these questions following their disability assessment interview with the claimant and after studying the information obtained from treating and/or occupational physician [36].

Prediction of Functioning by Disability Claimant At baseline, the prediction of improvement in functioning was assessed in claimants with a single question: “Do you expect improvement of your ability to function any time soon?”. Response categories were no or yes within 4 weeks, yes in between 4 weeks and 3 months, yes in between 3 and 6 months, yes in between 6 months and 1 year, and yes after 1 year. Answers were dichotomized into predicting improvements in functioning within 1 year versus after 1 year.

Improvement of Functioning

Functioning was measured at T0 and T1 with the 36-item self-reported version of the World Health Organization Disability Schedule 2.0 (WHODAS 2.0) [37]. The WHODAS 2.0 is a generic instrument that assesses levels of functioning during the previous 30 days in six domains of life: *Understanding and Communicating* (6 items), *Getting around* (5 items), *Self-care* (4 items), *Getting along with people* (5 items), *Life activities* (household activities: 4 items; work: 4 items) and *Participation in society* (8 items) [1]. Answering options are ‘none’ (1), ‘mild’ (2), ‘moderate’ (3), ‘severe’ (4) and ‘extreme/cannot do’ (5). The WHODAS 2.0 has high internal consistency (Cronbach’s alpha 0.86), a stable factor structure, high test–retest reliability (ICC: 0.98), good concurrent validity and good sensitivity to change [36]. For this study, the WHODAS scores excluding the work items were used as most disability claimants did not work. Scores were recoded and standardized using a SPSS syntax available on request from the WHO. Standardized total score and subscale scores range from 0 to 100 with higher scores representing increased difficulties in functioning. Standardized total scores were imputed by the mean if less than 10 % of the total scores were missing.

Analysis

Agreement between the predictions of IPs and claimants was calculated using the Kappa statistic. A kappa of <0.20 represents low agreement and a kappa of >0.60 is interpreted as high agreement [38]. Diagnostic accuracy of the predictions of IPs and disability claimants was assessed by calculating the sensitivity, specificity, and area under the receiver operating curves (AUC) with 95 % confidence intervals. The Receiver Operating Curve represents plots of the sensitivity and 1-specificity of the prediction of the improvements of functioning after 1 year. Moreover, Negative and Positive Predictive values (NPV and PPV) and Likelihood Ratios for positive and negative predictions (LR+ and LR−) were computed.

The prediction of improvement of functioning within 1 year (yes/no) was compared to the actual occurrence of relevant improvement of functioning 1 year later (yes/no). Improvement of functioning was defined as lower WHODAS 2.0 scores at T1 compared to T0. Only differences greater than the standard error of measurement (SEM) were considered relevant improvements as this provides a first indication of meaningful change [39]. The SEM agreement statistic was calculated using the following formula [40]:

$$\text{SEM\_agreement} = \sigma * \sqrt{(1 - \text{ICC\_agreement})}.$$

The variance and ICC_agreement were generated using the SPSS Reliability command, after which the standard deviation was also calculated as the square root of the variance. The ICC_agreement was derived from the "ICC single measures" in the SPSS output.

A generalized linear mixed model (GLMM) was constructed to calculate the probability of improvement as predicted by either IPs or disability claimants while accounting for the multilevel structure of the data (disability claimants within IPs). This predicted probability was used for calculating the AUC. Differences in overall accuracy of the predictions between IPs and disability claimants will be examined by comparing the Area Under the Curve and confidence intervals. An AUC of 0.50–0.70 is usually considered poor for any diagnostic test [41, 42]. Sensitivity, specificity, NPV, PPV, LR+, and LR− of the predictions of the IP and the claimant will be presented but not statistically tested as these characteristics do not take the multilevel structure of the data into account.

Differences in the accuracy of predictions in claimants with mental or somatic health conditions were tested by conducting a mixed logistic regression with correctness of the prediction (yes/no) as the dependent and primary medical diagnosis (mental vs. somatic) as the independent variable.

Analyses were performed using the statistical package IBM SPSS Statistics 19 (Armonk, NY, 2010).

Table 1 Baseline characteristics of disability claimants at T0, N = 375

Characteristic of disability claimants	
Gender, male, N (%)	190 (51)
Age in years, mean (SD)	50 (9)
Educational level	
Low, N (%)	66 (18)
Middle, N (%)	255 (68)
High, N (%)	45 (12)
Missing, N (%)	9 (2)
Diagnosis	
Mental, N (%)	84 (22)
Somatic, N (%)	282 (75)
Missing, N (%)	9 (2)
Employed in paid work	
Yes, N (%)	37 (10)
No, N (%)	338 (90)
Claimant expectation improvement of functioning	
≤1 year, N (%)	248 (66)
>1 year, N (%)	70 (19)
Missing, N (%)	57 (15)
IP expects improvement	
≤1 year, N (%)	125 (33)
>1 year, N (%)	106 (28)
Unsure, N (%)	51 (14)
Missing, N (%)	93 (25)

IP Insurance physician

Results

The characteristics of the disability claimants at baseline are presented in Table 1. Sixteen IPs reported data on the disability claimants included in the study, with a range from three to 37 claimants. Half of the disability claimants were male, the average age was 50 years and the majority had a medium-level education. Three quarters of the study population had received a primary somatic diagnosis and ten percent was still in some form of paid employment.

Prediction of Functioning

A greater proportion of the disability claimants (66 %) expected an improvement of functioning within 1 year, as compared to the expectation of IPs (33 %). A similar difference in proportions was seen when the missings and unsure categories were disregarded. After excluding these categories, 78 % of the claimants and 54 % of the IPs expected an improvement in the next year. Agreement between claimants and IPs was low, as indicated by a kappa of −.324.

Improvement of Functioning

At baseline, the disability claimants reported a mean total score of 34 (SD 18). One year later (T1) the mean total score was 30 (SD 19). All subscales showed lower means at T1 compared to T0. These improvements in the mean scores were lowest in self care (1.0; SD 15.7) and highest in life activities (7.4; SD 31.1).

The SEM_agreement for the standardized total scores at baseline and T1 was found to be 9.3, based on a $\sqrt{(1\text{-ICC\_agreement})}$ of 0.55 and a standard deviation of 17.06. The number of claimants that showed an improvement on overall functioning that exceeded the SEM was 84 (32 % of the 263 disability claimants included in this analysis).

Accuracy of Predictions Disability Claimants and IPs

Table 2 presents the accuracy of the predictions of improvement in functioning by claimants and their IPs. The predictions of the claimants had a sensitivity of 0.72 and a specificity of 0.21. The predictions of their IPs showed a lower sensitivity (0.52), but a higher specificity (0.47). With fairly similar NPV values for claimants and IPs (0.61 and 0.62 respectively), IPs showed a slightly higher PPV (0.37) compared to claimants (0.30). The LR + was low for both but slightly higher for IPs (0.99) compared to claimants (0.91). The LR- was 1.34 for claimants and 1.01 for IPs. The AUC were fairly similar between claimants and IPs (0.61 and 0.62 respectively; Fig. 2a, b). Both AUC were significantly higher than 0.5, pointing to a better prediction than one would have based on chance alone. The confidence intervals of the AUC of IPs and claimants showed great overlap.

Differences in Claimants with Somatic or Mental Conditions

Two separate logistic regressions were performed to examine differences in the accuracy of the predictions when claimants with mental or with somatic diagnoses were concerned. One model tested the relationship of diagnosis with correctness of the prediction of IPs. Of the 163 cases with full data, IPs predicted 80 correct (49 %) and a 83 cases incorrect (51 %). These 163 cases did not show statistically significant differences with regard to gender (45 % male), mean age (50.5; SD 8.7), and diagnosis (23 % mental) compared to cases with missing data (55 % male; mean age 49.2 (SD 9.1), and diagnosis (23 % mental). Predictions on claimants with a somatic diagnosis were not statistically better than predictions on claimants with mental diagnoses (OR 1.7; CI 0.8–3.6). The second model concerned the self predictions of 249 claimants. A hundred and fifty-five of these claimants provided incorrect predictions and 94 predicted their future status correctly. Claimants with a somatic diagnosis were not statistically better in predicting their status than claimants with mental diagnoses (OR 1.2; CI 0.63–2.1).

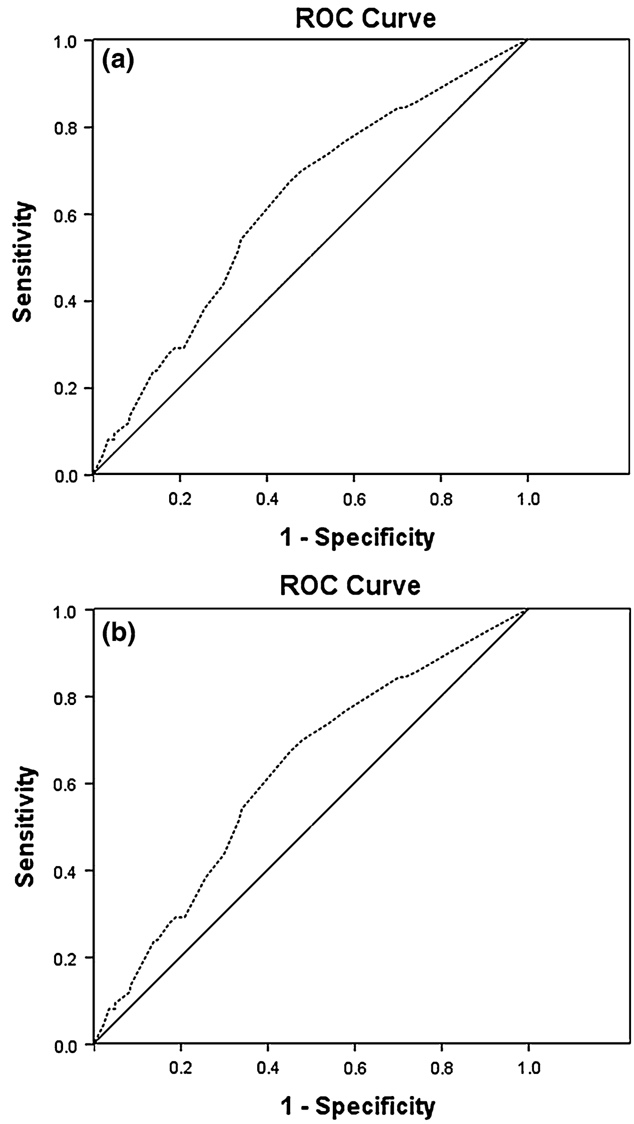


Fig. 2 Area under the curve of the predictions of claimants (**a**) and IPs (**b**)

Table 2 Accuracy of predictions of improvement of functioning by social insurance physicians (IPs) and disability claimants

	Functioning		Sensitivity	Specificity	AUC	(95 % CI)	PPV	NPV	LR+	LR−
	Improved	Not improved								
Prediction by claimant (data for N = 251)										
Improved	58	134	0.72	0.21	0.62[b]	(0.54–0.70)	0.30	0.61	0.91	1.34
Not improved	23	36								
Prediction by IP (full data for N = 163)[a]										
Improved	32	54	0.52	0.47	0.61	(0.53–0.69)	0.37	0.62	0.99	1.01
Not improved	29	48								

[a] Excluding the unsure category; *AUC* area under the curve, *PPV* positive predictive value, *NPV* negative predictive value, *LR+* likelihood ratio for a positive prediction; *LR−* likelihood ratio for a negative prediction

[b] The N for calculating the AUC was 220 as 31 claimants did not have participating IPs. These cases were excluded in the multilevel analysis as they could not be assigned to an IP

Discussion

This study showed that a substantial proportion of those claiming disability benefit after 2 years of sick leave show improvements of functioning the next year. Disability claimants' own prediction of future functioning was not more accurate than the predictions of the IP. Both were only able to predict future functioning with low levels of accuracy; an AUC of 0.61 for IPs and 0.62 for disability claimants. While differences were not statistically tested, claimants predicted improvements in a higher proportion of those who turned out to improve (higher sensitivity) and IPs predicted a lack of improvement in a higher proportions of those turning out not to improve (higher specificity). In the context of disability benefits, high sensitivity may prevent unnecessary long durations of disability benefits, while high specificity will be useful in preventing unnecessary re-assessments. IPS and claimants did not seem to differ in the overall accuracy of their predictions. We further found that the accuracy of the predictions did not show a statistically significant relationship with the diagnosis of the claimant (mental or somatic health conditions).

The proportion of accurate predictions by IPs found in our study was 49 %. That is fairly comparable to the 53 % of accurate predictions of absence status in workers on sick leave found in a study with general practitioners [15]. The general practitioners were asked to predict a status 4 weeks later, while the IPs in our study were asked to predict a status 1 year later. When evaluating the diagnostic value of predictions by IPs, one should preferably look at the AUC outcomes as these take the multilevel structure of the data into account. The AUC of the IP predictions did show that they predicted better compared to chance (AUC 0.50), but should still be considered poor diagnostic accuracy. However, usually diagnostic accuracy refers to the ability of a test to detect a health condition in the present time. In our study, changes over 1 year were the object of the predictions, which are harder to predict.

In contrast to what was found in earlier studies comparing professionals to self-predictions [26], the self predictions of disability claimants were not more accurate than the IP's assessments. As our measure of functioning is based on self-report, higher accuracy of self-predictions was expected. It should further be noted that IPs conducting the disability assessment had a face-to-face interview with the claimants which in theory may have led to more agreement in the predictions of IP and claimant due to discussing the health condition and functional limitations. Whether IPs discussed their expectations about future improvements in functioning with the claimant during the disability assessment is not known.

There are some aspects of our study that deserve further discussion. As our study examines the accuracy of predictions of future functioning, the definition of improvement of functioning deserves some consideration. First, the instrument used as a gold standard of functioning was the WHODAS 2.0 which measures functioning in six domains of life. An advantage of this instrument is that it can be used in disability claimants with various health conditions as it does not measure disease-specific limitations in functioning. The instrument is also suitable for claimants who are no longer in paid employment as only a small part refers to work activities or participation while the subscales reflect preconditions for being able to work. The downside of such a broad instrument may be that it may not reflect what disability claimants have in mind when asked about their functioning. Disability claimants and IPs may take the context of the specific health condition and of work opportunities into account when making their predictions, leading to less accurate predictions when compared to the WHODAS 2.0 total scale. Posthoc analyses revealed that the accuracy for predictions based on improvements on the life activities subscale alone, the one subscale most closely related to work functioning, was similar as for the total scale (AUC of 0.62 for claimants and IPs).

Also the assessments of IP and claimant predictions were conducted with slightly different answer categories (ordinal versus ratio scale, both dichotomized in the analysis). Studies assessing predictions from professionals or self-predictions have used different answer categories, see e.g. [15, 19, 26]. How the answer category is linked to the accuracy of the predictions has, to our knowledge, not yet been studied.

A further discussion point related to our gold standard is the definition of improvement as an improvement greater than the SEM (9.3 points). This criterion reflects some level of relevancy, since improvements that are likely to be due to measurement error are disregarded. However, we do not know whether the improvements observed in our sample reflect important changes. It would be preferable to compare changes in functioning to a value of the minimal important change (MIC) [43]. The MIC value refers to the smallest difference between two scores on a measurement scale that can be regarded as relevant or important [44]. Future studies are needed to define what disability claimants and IPs consider an important change in functioning in order to calculate the MIC of instruments of overall functioning.

The PREDIS cohort study had a low response rate, which may have been caused by the burden of the clinical interview that was part of this study and lasted approximately 2–4 h, depending on the mental health status. The representativeness of the participants could only be tested in relation to the primary cause for disability. No statistical differences were found between participants and the overall SSI population. Moreover, full data were available in a lower number of cases, especially concerning the accuracy of IP predictions. This is partly due to IPs answering "unsure" or not at all in 40 % of the cases. It is unlikely that IPs being unsure or not wanting to answer who would have predicted future functioning accurately. The diagnostic accuracy found in this study should therefore probably be considered an upper limit of accuracy.

A further limitation of this study was that it used the primary ICD-10 diagnosis to distinguish between claimants with mental or with somatic diagnoses. These diagnoses are certified by the IPs based on information in the medical files and their disability assessment. Prior studies with the PREDIS cohort have shown that many claimants classified as having a somatic diagnosis also have mental disorders according to a structured clinical interview based on the DSM-IV classification [Cornelius et al. submitted]. We have therefore conducted a post hoc analysis to check whether the accuracy of the IP prediction was worse in claimants who had a mental co-morbidity. A logistic regression revealed that the prediction in this group was not less often correct compared to claimants with primary mental disorders or somatic disorders without mental co-morbidity (OR 0.86; CI 0.41–1.8).

The accuracy of predictions of improvement of functioning by both IPs and claimants was not high. It appears that selecting claimants eligible for re-assessments should not be based on these predictions alone. One may argue that the time period between the predictions and the predicted improvements 1 year later may be too long to allow for high levels of accuracy. However, regular re-assessments by IPs may not be feasible. An alternative that may be considered is to ask claimants to periodically fill out self-report instruments such as the WHODAS 2.0 to monitor functioning over time in the period after the first disability assessment. However, before one can use self report instruments to monitor improvements, the minimal important change values need to be established. In the context of disability assessment, additional requirements need to be met. In that context important improvements should not only be the claimant's view, but should also reflect an improvement likely to affect the outcome of a disability assessment by an IP.

Predicting future functioning of disability claimants remains a challenge. The IPs in this study were provided with information obtained from treating and/or occupational physicians and conducted disability assessment interviews with each of the patients. Nevertheless, predictability of functioning may be enhanced by more emphasis on non-medical aspects of the prognosis. A recent Delphi study among IPs concluded that non-medical personal and environmental factors must be considered in the assessment of the work ability of long-term sick-listed employees [45].

Moreover, future studies on the PREDIS cohort will be conducted aiming to identify other predictors of functional improvement and work status after the disability claim assessment. These predictors, if identified, can be used in a prognostic tool to assess eligibility for re-assessment.

In conclusion, we found that improvements of functioning are not uncommon in a group of disability claimants with various diagnoses. Both IPs and disability claimants themselves were only able to predict improvements with low levels of accuracy. Alternative ways to screen for eligibility for re-assessments, such as the use of self report instruments over time, should be considered.

Acknowledgments This research project was funded by the Social Security Institute, the Netherlands. The funding institute had no role in the design, collection, analysis, and interpretation of data; in the writing of the manuscript; and in the decision to submit the manuscript for publication.

References

1. Organisation for Economic Co-operation and Development (OECD). Sickness, disability and work: keeping on track in the economic downturn. Background paper. http://www.oecd.org/employment/employmentpoliciesanddata/42699911.pdf. Accessed 10 Aug 2012.
2. Organisation for Economic Co-operation and Development (OECD). Sickness, disability and work: braking the barriers. http://www.nho.no/getfile.php/bilder/RootNY/filer_og_vedlegg1/Synthesis%20OECD%20Sickness.pdf. Accessed 20 Dec 2012.
3. Chorus AM, Miedema HS, Boonen A, et al. Quality of life and work in patients with rheumatoid arthritis and ankylosing spondylitis of working age. Ann Rheum Dis. 2003;62(12):1178–84.
4. Main DS, Nowels CT, Cavender TA, et al. A qualitative study of work and work return in cancer survivors. Psychooncology. 2005;14(11):992–1004.
5. Dunn EC, Wewiorski NJ, Rogers ES. The meaning and importance of employment to people in recovery from serious mental illness: results of a qualitative study. Psychiatr Rehabil J. 2008;32(1):59–62.
6. Gronning K, Rodevand E, Steinsbekk A. Paid work is associated with improved health-related quality of life in patients with rheumatoid arthritis. Clin Rheumatol. 2010;29(11):1317–22.
7. Bernklev T, Jahnsen J, Henriksen M, et al. Relationship between sick leave, unemployment, disability, and health-related quality of life in patients with inflammatory bowel disease. Inflamm Bowel Dis. 2006;12(5):402–12.
8. Melchior M, Ferrie JE, Alexanderson K, et al. Does sickness absence due to psychiatric disorder predict cause-specific mortality? A 16-year follow-up of the GAZEL occupational cohort study. Am J Epidemiol. 2010;172(6):700–7.
9. Ferrie JE, Vahtera J, Kivimaki M, et al. Diagnosis-specific sickness absence and all-cause mortality in the GAZEL study. J Epidemiol Commun Health. 2009;63(1):50–5.
10. Lund T, Kivimaki M, Christensen KB, et al. Socio-economic differences in the association between sickness absence and mortality: the prospective DREAM study of Danish private sector employees. Occup Environ Med. 2009;66(3):150–3.
11. Richards D. Prevalence and clinical course of depression: a review. Clin Psychol Rev. 2011;31(7):1117–25.
12. Waxman R, Tennant A, Helliwell P. A prospective follow-up study of low back pain in the community. Spine (Phila Pa 1976). 2000;25(16):2085–90.

13. De Boer WEL, Besseling JJM, Willems JHBM. Organisation of disability evaluation in 15 countries. Pratiques et organisation des soins. 2007;38(3):205–17.
14. Stattin M. Retirement on grounds of ill health. Occup Environ Med. 2005;62(2):135–40.
15. Reiso H, Gulbrandsen P, Brage S. Doctors' prediction of certified sickness absence. Fam Pract. 2004;21(2):192–8.
16. Waddell G, Burton AK. Concepts of rehabilitation for the management of common health problems. Norwich: TSO (The Stationery Office); 2004. http://www.dwp.gov.uk/docs/hwwb-concepts-of-rehabilitation.pdf. Accessed at 14 Aug 2012.
17. Dekkers-Sanchez PM, Wind H, Sluiter JK, et al. A qualitative study of perpetuating factors for long term sick leave and promoting factors for return to work: chronic work disabled patients in their own words. J Rehabil Med. 2010;42(6):544–52.
18. Audhoe SS, Hoving JL, Nieuwenhuijsen K, et al. Prognostic factors for the work participation of sick-listed unemployed and temporary agency workers with psychological problems. J Occup Rehabil. 2012;22(4):437–46.
19. Nielsen MB, Madsen IE, Bultmann U, et al. Predictors of return to work in employees sick-listed with mental health problems: findings from a longitudinal study. Eur J Public Health. 2011;21(6):806–11.
20. Nieuwenhuijsen K, Verbeek JH, De Boer AG, et al. Predicting the duration of sickness absence for patients with common mental disorders in occupational health care. Scand J Work Environ Health. 2006;32(1):67–74.
21. Lagerveld SE, Blonk RWB, Brenninkmeijer V, et al. Return to work among employees with mental health problems: development and validation of a self-efficacy questionnaire. Work Stress. 2010;24(4):359–75.
22. Richard S, Dionne CE, Nouwen A. Self-efficacy and health locus of control: relationship to occupational disability among workers with back pain. J Occup Rehabil. 2011;21(3):421–30.
23. Shaw WS, Reme SE, Linton SJ, et al. 3rd place, PREMUS best paper competition: development of the return-to-work self-efficacy (RTWSE-19) questionnaire–psychometric properties and predictive validity. Scand J Work Environ Health. 2011;37(2):109–19.
24. Gross DP, Battie MC. Recovery expectations predict recovery in workers with back pain but not other musculoskeletal conditions. J Spinal Disord Tech. 2010;23(7):451–6.
25. Cole DC, Mondloch MV, Hogg-Johnson S. Listening to injured workers: how recovery expectations predict outcomes–a prospective study. Can Med Assoc J. 2002;166(6):749–54.
26. Fleten N, Johnsen R, Forde OH. Length of sick leave—why not ask the sick-listed? Sick-listed individuals predict their length of sick leave more accurately than professionals. BMC Public Health. 2004;4:46.
27. Wind H, Samoocha D, van der Beek AJ, et al. Prevention of disability: the opinion of claimants applying for a disability benefit. Work. 2013. doi:10.3233/WOR-131663.
28. Brouwer S, Reneman MF, Bultmann U, et al. A Prospective study of return to work across health conditions: perceived work attitude, self-efficacy and perceived social support. J Occup Rehabil. 2010;20:104–12.
29. Purdie F, Kellett S, Bickerstaffe D. Predictors of functional disability in disability welfare claimants. J Occup Rehabil. 2012;22:447–55.
30. Sampere M, Gimeno D, Serra C, et al. Return to Work expectations of workers on long-term non-work-related sick leave. J Occup Rehabil. 2012;22:15–26.
31. Knowledge Center UWV. Knowledge center UWV quarterly report 2007-III. Dutch: Kenniscentrum UWV Kwartaalverkenning 2007-III; 2009. http://www.uwv.nl. Accessed 15 Jan 2009.
32. UWV. Quantitative information 2010. Dutch: Kwantitatieve Informatie; 2010. http://www.uwv.nl/overuwv/Images/Kwantitatieve_informatie_2010.pdf. Accessed 11 July 2013.
33. Kessler RC, Ustun TB. The world mental health (WMH) survey initiative version of the world health organization (WHO) composite international diagnostic interview (CIDI). Int J Methods Psychiatr Res. 2004;13(2):93–121.
34. World Health Organization. International statistical classification of diseases and related health problems 10th revision (ICD-10); 2012. http://apps.who.int/classifications/icd10/browse/2010/en. Accessed 19 Apr 2012.
35. Ouwehand P, Wouters PHM. CAS, Classificaties voor Arbo en SV (CAS, Classification for Occupational Health and Social Security). Amsterdam: UWV; 2002. http://www.debus.nl/images/PDF/cas-code%20boekje.pdf. Accessed 15 Jan 2009.
36. Spanjer J, Krol B, Popping R, et al. Disability assessment interview: the role of detailed information on functioning in addition to medical history-taking. J Rehabil Med. 2009;41(4):267–72.
37. World Health Organization. Measuring health and disability: manual for WHO disability assessment schedule (WHODAS 2.0). In: Üstün TB, Ostanjsek N, Hatterji S, Rehm J (eds.). 2010. http://whqlibdoc.who.int/publications/2010/9789241547598_eng.pdf. Accessed 19 Apr 2012.
38. Altman D. Practical statistics for medical research. London: Chapman & Hall; 1991.
39. McLeod LD, Coon CD, Martin SA, et al. Interpreting patient-reported outcome results: US FDA guidance and emerging methods. Expert Rev Pharmacoecon Outcomes Res. 2011;11(2):163–9.
40. de Vet HC, Terwee CB, Knol DL, et al. When to use agreement versus reliability measures. J Clin Epidemiol. 2006;59(10):1033–9.
41. Swets JA. Measuring the accuracy of diagnostic systems. Science. 1988;240(4857):1285–93.
42. Fischer JE, Bachmann LM, Jaeschke R. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. Intensive Care Med. 2003;29(7):1043–51.
43. Terwee CB, Roorda LD, Dekker J, et al. Mind the MIC: large variation among populations and methods. J Clin Epidemiol. 2010;63(5):524–34.
44. Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. Curr Opin Rheumatol. 2002;14(2):109–14.
45. Dekkers-Sanchez PM, Wind H, Sluiter JK, et al. What factors are most relevant to the assessment of work ability of employees on long-term sick leave? The physicians' perspective. Int Arch Occup Environ Health. 2013;86(5):509–18.